DOCUMENT RESUME

ED 064 360 TM 001 528

AUTHOR Greenberg, Barry; Mejias, Ricardo
TITLE A Study of the Use of Multiple Regression with Dummy Variables to Identify Instructor Contribution to Student Achievement.
PUB DATE Apr 72
NOTE 29p.; Paper presented at the annual meeting of the AERA (Chicago, Ill., April 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Academic Achievement; *Educational Accountability; Evaluation Techniques; *Multiple Regression Analysis; Student Evaluation; *Teacher Evaluation; Teacher Role; *Test Results

ABSTRACT

This study utilized linear multiple regression analysis with dummy variables to isolate one component of an educational accountability system, the contribution of the individual teacher to student achievement. Independent variables consisted of measures of each student's past academic record, the size of the class and a dummy variable to represent the student's instructor. The dependent variable was scores on a Social Science departmental final examination. Results revealed the effectiveness of this approach in isolating and rank ordering each instructor's contribution to student achievement. Obtained regression coefficients for instructors paralleled department chairman perceptions of their contribution. The discussion considered the possibility of this method for faculty evaluation and included some recommendations for those contemplating its use (Author)

A STUDY OF THE USE OF MULTIPLE REGRESSION WITH DUMMY VARIABLES TO IDENTIFY INSTRUCTOR CONTRIBUTION TO STUDENT ACHIEVEMENT

Barry Greenberg and Ricardo Mejias
Miami-Dade Junior College
South Campus

Paper presented at the annual meeting of the
American Educational Research Association,
Chicago, April 4, 1972.

Abstract

This study utilized linear multiple regression analysis with dummy variables to isolate one component of an educational accountability system, the contribution of the individual teacher to student achievement. Independent variables consisted of measures of each student's past academic record, the size of the class and a dummy variable to represent the student's instructor. The dependent variable was scores on a Social Science departmental final examination. Results revealed the effectiveness of this approach in isolating and rank ordering each instructor's contribution to student achievement. Obtained regression coefficients for instructors paralleled department chairman perceptions of their contribution. The discussion considered the possibi ty of this method for faculty evaluation and included some recommendations for those contemplating its use.

INTRODUCTION

As accountability in public education becomes required practice, it is expected that one of the more fundamental shifts it will bring about will be in the area of faculty evaluation. Demands for "cost-effective" education presume the ability to quantify the individual teacher's contribution to the learning enterprise. That faculty evaluation in the past has not been synonymous with quantification of the teacher's contribution to pupil learning hardly needs elaboration. The former has dwelt primarily upon credit hours completed, years of experience and an occasional supervisory visit. For the most part, each of these criteria have proved to be only marginally related to student outcomes or any other meaningful indicator of school quality. Indeed, in a recent review of the literature of faculty evaluation, Cohen and Brawer concluded that " ..... the best that can be said for current methods for evaluation of faculty in institutions of higher education is that they are ineffectual and little regarded." (1969, pg.2).

This state of affairs can hardly continue in the midst of movement toward systems of program budgeting with its emphasis on outputs and costs. Since the largest item, by far, in all school budgets is for

instructional salaries, it is expected that the need for objective indices of teacher effectiveness will be more urgently felt than ever before. The purpose of this study was to investigate one approach toward capturing this needed data.

Related Literature

For some time now, disenchantment with "process" evaluation procedures has been growing. Major points of critics have been that processes (or, what teachers do or come to the classroom with) are not of themselves part of the reasons for the existence of schools. They are used as indicators, or, "proxies", of true educational outputs. Thus, it has been assumed that the teacher who utilizes a certain questionning technique for example, is more effective than one who does not. The investigation typically ceased with the discovery of the presence of absence of the questionning process. It served as a proxy for teacher effectiveness. Many have pointed out that this tendency has been dysfunctional to the accomplishment of the objectives of education.

Methods which look at product or output, in short what students learn rather than the process of what teachers do, have been suggested as a realistic alternative to process evaluation. The research herein reported derived from developmental work of Barro (1970) and findings of Hanushek (1970) both of whom recommended a

linear multiple regression approach to isolate the contribution of the individual teacher to the achievement of his students. In utilizing regression with dummy variables Hanushek was able to capture and test for statistical significance the individual teacher's contribution to student achievement. Previous research of Astin and Panos (1969) has demonstrated the utility of a multiple regression approach. For these authors, multiple regression proved useful in assessing the effectiveness of 246 institutions of higher education. The technique enabled them to satistically "equate" the entering freshmen at each school and then compare outputs. That school whose actual output exceeded its predicted output (based on the characteristics of its entering students) might then be thought of as more successful than an institution which produced the reverse effect (e.g., actual output below predicted output).

More recently, Astin (1972) utilized multiple regression to identify student characteristics associated with dropping out of college and to identify the colleges most effective in this area, and, Centra and Rock (1971) utilized multiple regression analysis to determine the influence of college environmental factors on student achievement controlled for entering differences.

Barro noted that adding what are called dummy variables or "identifiers" to the multiple regression analysis "would permit direct estimation of the degree to which pupil performance in each classroom differs from pupil performance in the average classroom." (cf. 1970, pg. 3) This approach creates a "norm"

for each classroom based on the characteristics of the enrolled students. If the norm is exceeded, the coefficient of the dummy variable is positive. If the teacher has not been able to equal his "norm" his coefficient will be negative. In utilizing this approach Hanushek was able to derive coefficients for the contributions of teachers to the achievement of second and third graders.

Population and Methodology

The South Campus of Miami-Dade Junior College served as the setting for this study. The sample was composed of nine instructors who taught 36 sections of Social Science 102 in the Winter 1970 semester. Social Science 102 is a required general education course in economics and political science. The number of students used in the analysis, n=572, was limited to those for whom input data existed.

The method utilized was linear least-squares multiple regression analysis. The dependent variable was the student's raw score on the Social Science final examination which is administered to all students in all sections who complete the course.

The independent variables consisted of:

1. Florida State-Wide Twelfth Grade Test Percentiles (F.T.G.) in:
   a. English
   b. Aptitude
   c. Social Science

These tests are administered to all Florida students in their senior year of high school.

2. Grade Point Averages including the Winter, 1970 semester but excluding any grade earned in Social Science 102.
3. Class Size, defined as the number of students who took the final examination.
4. Cumulative Hours Earned, including the Winter, 1970 semester but excluding credit for Social Science 102.
5. 9 Instructor Dummy Variables, one for each instructor.

The dummy variables provided the mechanism to determine whether students "equalized" on the input variables (items 1-4 above) were able to achieve more -- defined as scoring higher on the final examination -- with one instructor than with another.

Additional data obtained were the department chairman's perception of the contribution made to student achievement by each of the nine instructors of the study, the number of years of teaching experience for each instructor and the number of credit hours earned beyond the Master's degree by each instructor.

The methodology was appropriate for the following objectives:

1. To derive instructor effectiveness coefficients which control for differences in certain student characteristics.
2. To identify the contribution to student achievement made by available student characteristics and class size.
3. To discover what relationships, if any, exist between the derived instructor effectiveness coefficients and:
   a. their department chairman's perception of their contribution to student achievement.
   b. their years of teaching experience.
   c. the number of credits they have earned beyond the Master's degree.

The first objective focused on providing data for one component of an accountability system, the contribution of the individual teacher to student achievement.

Objective two enabled comparisons to be made amongst independent variables insofar as their significance for achievement was concerned. Objective three was, in a sense, a validity check enabling comparisons to be made between the department chairman's perceptions of teacher contribution and the traditional pay parameters (credits and years of experience) with the dummy variable coefficients herein discussed.

RESULTS

1. Relationship Between the Students' Academic Background and Scores on the Social Science 102 Final Examination.

Table 1 presents the results of regressing the final examination scores on five variables of student academic background, class size and nine instructor dummy variables. Among the measures of academic background, the Florida Twelfth Grade Social Science Test (FTGSSC) was the best predictor of performance in the SSC 102 final examination. Grade point average (GPA) and cumulative hours earned in college (CHE) were also significant predictors, while, FTGAPT and FTGENG show no significant relationship to final examination scores. Nevertheless, other regressions, with FTGSSC excluded, show the latter two to be statistically significant at the 1% level, but less significant than FTGSSC. Furthermore, the zero-order correlation coefficient between FTGSSC and FTGAPT is 0.64, and between FTGSSC and FTGENG, it is 0.46. This means that multicollinearity (correlation among the independent variables) makes it impossible to separate the individual association of each FTG test with the final examination score. But, the insignificance of FTGAPT and FTGENG when FTGSSC is included, means that there is no additional information that these tests can provide towards predicting final examination scores beyond what is provided by the FTGSSC percentiles.

TABLE 1

Regression of achievement on all student input variables, class size and instructors.

Equation (1)

Dependent Variable:			
SSCF			
Independent Variables:	Coefficients:	F Values	
Constant	18.636		
FTGAPT	0.012	0.284	$\bar{R}^2 = 0.47$
FTGENG	-0.002	0.001	
FTGSSC	0.191	106.788	
CHE	0.113	7.838	Critical F:
GPA	5.923	90.398	
CL SIZE	-0.052	1.121	at 5% level: 1.80
INST 1	-6.098	8.696	
INST 2	-1.430	0.770	at 1% level: 2.28
INST 3	2.206	0.880	
INST 4	-2.335	2.479	
INST 5	0.119	0.004	
INST 6	-1.035	0.505	
INST 7	-1.312	0.778	
INST 8	1.548	0.483	
INST 9	0.000	0.000	

For simplification, FTGAPT and FTGENG were deleted from the second regression presented as Table 2. In the Table, class size has been excluded from equation (2), but included in equation (3), for reasons explained below. Both CHE (cumulative hours earned in college) and GPA, as well as FTGSSC, are related to the SSC 102 final at the 1% level of significance. As could be expected, there is some multicollinearity, as shown on Table 3. Consequently, the individual association between each of these variables and the SSC 102 scores cannot be fully isolated. But, this multicollinearity is relatively low, and since the "F" values are so high, there appears to be a considerable degree of independent association between each of these variables and the SSC 102 final, in addition to their joint association. Nevertheless, the FTGSSC test, taken 1½ years before the SSC 102 final examination was given, explains more of the variation in final examination scores than GPA, CHE, or any other variable included.

The above point to the importance of measures of academic background that are specific to the subject matter, as predictors of performance on that subject. Consequently, it is very likely that a pre-test, specifically designed for the content of the course, would improve upon the ability to adjust for student academic background, in attempting to isolate the instructor's contribution to student achievement. The results also indicate that a measure of general academic background, such as GPA, is also very helpful for this purpose.

TABLE 2

Regression of achievement on student ability (excluding FTGAPT and FTGENG), class size (only in equation 3) and instructors.

Dependent Variable:	(2)		(3)	
SSCF				
Independent Variables:	$\bar{R}^2 = 0.47$ Coefficients (F Value)	Instructor Rank	$\bar{R}^2 = 0.47$ Coefficients (F Value)	Instructor Rank
Constant	11.667		12.583	
FTGSSC	0.197 (177.392)		0.197 (177.181)	
CHE	0.113 (7.999)		0.115 (8.164)	
GPA	5.991 (104.051)		5.971 (103.246)	
CLSIZE			-0.052* (1.114)	
INST 1	0.000 (0.000)	9th	0.000 (0.000)	9th
INST 2	3.663 (4.284)	7th	4.673 (5.395)	7th
INST 3	7.056 (8.353)	1st	8.285 (9.383)	1st
INST 4	2.546 (2.479)	8th	3.791 (3.588)	8th
INST 5	5.858 (10.544)	2nd	6.265 (11.536)	3rd
INST 6	4.424 (7.836)	5th	5.067 (8.951)	5th
INST 7	3.815 (5.405)	6th	4.794 (6.469)	6th
INST 8	5.841 (7.304)	3rd	7.739 (7.578)	2nd
INST 9	5.142 (7.838)	4th	6.135 (8.640)	4th

Critical F:
at 5% level: 1.80
at 1% level: 2.28

*Significant at the 30% level.

TABLE 3

Zero-order Correlations among FTG SSC, CHE, GPA and CL SIZE

	FTGSSC	CHE	GPA	CL SIZE
FTGSSC	1.00	0.11	0.33	-0.02
CHE		1.00	0.34	0.02
GPA			1.00	-0.02
CL SIZE				1.00

2. The Relationship Between Class Size and Scores on the SSC 102 Final Examination.

The issue of the effect of class size on learning has been a very controversial one. For long, it has been popularly believed that, the larger the class size, the less each student will learn. On the other hand, the overwhelming majority of educational research studies have failed to uncover any relationship between class size and learning, however measured.

Class size was found to be negatively, but not significantly related to scores on the SSC 102 final examination. (See equation (3), Table 2). Based on the "F" test, the probability that the relationship was obtained by chance is 30%.

There are reasons to believe however, that in this study the relationship may have been weakened by "simultaneous equation bias." That is, it appears that instructors who tend to contribute more to their students' final examination scores were assigned larger classes. As a result, the effect of class size may not have been completely isolated from the effect of

the instructor.

3. Differences Among Instructors in their Contribution to Student Achievement in the Social Science 102 Final Examination.

The dummy variable coefficients of equations (2) and (3) measure the differences in final examination scores attributable to each instructor, as compared to instructor 1, independent of the academic ability of the instructors' students and (in equation (3) only) independent of class size. But this statement is only correct to the extent that the independent variables used actually measure academic ability, and to the extent that the effect of class size was correctly accounted for. This does not mean that every single factor that affects a student's final examination score must be known for this method to correctly isolate the instructor's contribution. For this, it is only required that the factors that were not considered affect the final examination scores equally among all the instructors. For example, if the sex of the student also independently affects the final examination score, introducing sex as another independent variable would increase the explained portion of the variations in final examination scores (would raise $\bar{R}^{-2}$), but it would not change the values of the instructor dummy variable coefficients unless some instructors had a greater proportion of females than others.

As can be seen on Table 2, whether or not class size is used as an independent variable, affects the value of the instructor coefficients, and alters the rank order of instructors

5 and 8. This is an important reason why further research on the class size issue is needed before this method could be used for any actual evaluative purposes. But for purposes of analyzing the results of instructor contribution to student achievement, equation (3) has been used, with class size included, because the statistical probability that class size is associated with student achievement (70%) is greater than the probability than it is not (30%).

Since the instructor coefficients indicate the differences in examination scores between each instructor and instructor 1 which are attributable to the instructor, what the "F" test of Table 2 shows is that all instructors have made a contribution to their students' final examination scores which is significantly greater than that of instructor 1, at the 1% level. In order to compute the significance of the difference between each instructor and every other instructor, each of the two equations on Table 2 was run nine times with each of the nine instructors, in turn,becoming the standard of comparison. From this information Table 4 was compiled, which shows, for each instructor, which other instructors had dummy variable coefficients which were statistically different from his at the 5% level.

To evaluate the differences between the comparisons of equation (3), and comparisons based on unadjusted final examination scores, the reader is referred to Table 5. It can be seen that there are substantial differences between

the results of each method. Therefore, failing to adjust for differences in the student and classroom characteristics, can lead to attributing to an instructor either more, or less teaching effectiveness than would be attributed when these are considered. In the case of instructor 2, the difference was quite large. His rank order was changed from 2nd to 7th when the portion of the difference in the final examination scores attributable to student characteristics and class size was isolated from the portion attributable to the instructor. These results lead one to question the validity of any faculty evaluation process that does not control for differences in the student's entering academic ability (and possibly class size).

TABLE 4

STATISTICAL SIGNIFICANCE OF THE DIFFERENCE IN DUMMY VARIABLE COEFFICIENTS BETWEEN INSTRUCTORS IN EQUATIONS 2 AND 3 OF TABLE 2.

(2)		(3)	
Instructor Number:	Significantly Different from: (at 5% level)	Instructor Number:	Significantly Different from: (at 5% level)
1	2-9	1	2-9
2	1, 2, 5	2	1, 3, 8
3	1, 2, 3, 7	3	1, 2, 4, 6, 7
4	1, 3, 5, 8, 9	4	1, 3, 5, 8, 9
5	1, 2, 4, 7	5	1, 3
6	1, 4	6	1, 3
7	1, 2, 5	7	1, 3, 8
8	1, 4	8	1, 2, 4, 7
9	1, 4	9	1, 4

TABLE 5

COMPARISON OF INSTRUCTOR DUMMY VARIABLE COEFFICIENTS ON EQUATION (3) WITH DIFFERENCES IN MEANS OF FINAL EXAMINATION SCORES BETWEEN THE 9 INSTRUCTORS

(1)	(2)	(3)	(4)	(5)	(6)
Instructor Number	Final Means	Final Mean Differences From Inst. 1	Instructor Rank	Dummy Coefficients	Instructor Rank
1	44.33	0.00	7th	0.00	9th
2	46.65	2.32	2nd	4.67	7th
3	50.88	6.55	1st	8.28	1st
4	42.78	-1.55	9th	3.79	8th
5	45.65	1.32	5th	6.26	3rd
6	45.71	1.38	4th	5.07	5th
7	43.95	-0.38	8th	4.79	6th
8	46.30	1.97	3rd	7.74	2nd
9	45.59	1.25	6th	6.13	4th

4. <u>The Instructors' Contribution to the Final Examination Scores Compared to Measures of Faculty Evaluation and to Salaries.</u>

At Miami-Dade Junior College, considerable stress is placed on the importance of the teaching function of the faculty. With the Social Institutions Department having a long tradition of giving common departmental final examinations, it was of interest to compare the dummy variable coefficients with other forms of faculty evaluation.

a. The Department Chairman's Assessment.

The Social Institutions Department changed chairmen eight months prior to the time when the final examination herein analyzed was given. Since the first chairman headed this department for the previous five years and knew its faculty well, it was decided to include him, as well as the second chairman, in this study.

Each chairman was asked to do the following: "Rate these nine instructors in order of how much you think they contribute to the students' learning, as measured by their final examination scores, independently of any other qualities on which instructors could be evaluated." At the time of the interview, neither chairman had seen any of the results of this study.

The first chairman had headed the Social Institutions Department since the opening of the South Campus in 1964. He was assigned to chair another department at the end of the 1968-69 academic year, and he was interviewed in July of 1971.

He stated that his evaluation was besed on classroom visitations, and that he had never made comparisons of average final examination scores between the instructors.

The second chairman headed the department between the Fall Term of 1969 and the Summer Term of 1971, when he was assigned to chair another department. He was interviewed in November of 1971, and stated that his evaluation was based on personal conversations with the instructors, and that he had also made classroom visitations. He also stated that he had previously seen the final examination average scores of each instructor, but that he could not remember anything about them at the time of the interview. The second chairman's evaluation only included eight of the nine instructors because he was unable to get to know one of them.

To compare the chairman's evaluation to the earlier results of this paper, the instructors' dummy variable coefficients were regressed on the rank ordering of instructors provided by the chairmen. For the first chairman, the regression was significant at the 5% level, and close to the 1% level.

(4) $IDVC = 1.78 + 0.68\ CHRK_1$ $\quad \bar{R}^2 = 0.53$
Computed F = 10.07

5% critical F = 5.32
1% critical F = 11.26

IDVC: Instructors dummy variable coefficients.
$CHRK_1$: First chairman's ranking of instructors.

In the case of the second chairman, the regression was too insignificant to be worth computing, since the $\bar{R}^2$ was only 0.0006. To attempt to detect whether there was any relationship

between unadjusted differences in averages of final examination scores, and the chairman's evaluation of the instructors, these were also regressed. The equation for the first chairman was very close to being significant at the 5% level, but its significance and the $\bar{R}^2$ were well below those of equation (4), where the instructors dummy variables constitute the dependent variable.

(5) $FMD = -1.27 + 0.54\ CHRK_1$ $\qquad \bar{R}^2 = 0.34$
Computed F = 5.10

5% critical F = 5.32

FMD: Differences in final examination means.
$CHRK_1$: First Chairman's ranking of instructors.

In the case of the second chairman, the regression between his ranking of instructors and differences in final examination means was further from being significant at the 5% level, as compared to equation (5), but, this equation had a better fit ($\bar{R}^2 = 0.28$) than the one regressing the instructors dummy variable coefficients on the second chairman's evaluation.

(6) $FMD = -0.62 + 0.53\ CHRK_2$ $\qquad \bar{R}^2 = 0.28$
Computed F = 3.74

5% critical F = 5.59

It would not be appropriate to derive any general conclusions about the way in which chairmen evaluate their faculty from a sample of only 2 chairmen and 9 instructors. What has been presented in this section is only an illustration of the possibilities that the multiple regression method has as a tool for investigating relationships between the teaching and the

evaluation processes. Aside from this, the very close relationship that was found between the instructors' dummy variable coefficients and one of the two chairmen's evaluations strengthens confidence in the regression results. Both estimations of instructors teaching effectiveness were arrived at by totally different methods, and independently of each other. While their close agreement does not prove either of them correct, the finding of such remarkable correspondence would indicate that this subject merits some attention in the form of further research.

b. Years of Experience, Graduate Credits, Chairman's Evaluation and Salaries.

Years of experience and graduate credits beyond a Master's degree are used as major criteria for salary and rank at Miami-Dade Junior College. That these criteria were in fact applied to the nine instructors in this study is demonstrated by regressing their yearly salaries against these two variables:

(7) SALR = 8,294 + 199 TECHE + 33 CRDM
Computed F = 14.46 3.76 $\bar{R}^2 = 0.70$

5% critical F = 4.46
1% critical F = 8.65

r
TECHE, CRDM = -0.07

TECHE : Years of teaching experience.
CRDM: Credits beyond the Master's degree.

The relationship between salary and years of experience is significant at the 1% level, but that for credits beyond the Master's degree and salary does not quite reach the 5% level. Further, years of experience and credits beyond the Master's degree are truly independent variables, since the correlation coefficient between them is an insignificant -0.07.

The question of interest in this context was whether or not these salary criteria were related to the instructors' contribution to their students' final examination scores as herein attained. This was explored by regressing the instructor dummy variable coefficients on years of experience and credits beyond the Master's degree, as follows:

(8) IDVC = 7.82 - 0.31 TECHE - 0.05 CRDM
Computed F = 0.81 1.09 $\bar{R}^2 = 0.23$

IDVC = Instructors' dummy variable coefficients.
TECHE= Years of teaching experience.
CRDM = Credits beyond the Master's degree.

Both salary criteria variables are negatively related to the instructor dummy variable coefficients, but this relationship is far from being significant at the 5% level. The absence of any relationship between years of experience and the instructors' contribution to the final examination score is the most interesting one, since the years of experience criteria is usually justified as a proxy measure for teaching ability.

A final regression of instructor dummy variable coefficients on instructor salaries was also statistically insignificant. It may be therefore that years of teaching experience carries a weight in determining salaries that is well out of porportion to any importance it may have in determining teaching effectiveness.

Hanushek (1970), in a study of 2nd and 3rd grade reading achievement in a large California school district, also found no relationship between graduate credits, years of teaching

experience, and the instructor's contribution to student achievement. (He did find the instructor's contribution to be significantly related to his score on a verbal facility test, and to the number of years since his most recent educational experience.)

The importance of finding what instructor characteristics do contribute to student achievement of educational objectives cannot be overemphasized. As Hanushek pointed out, schools seem to be paying for the wrong kind of features in their teachers. Schools that can more correctly identify the characteristics of instructors that lead to better achievement of their objectives should be able to make considerable gains in achieving these objectives. These gains can be obtained in two ways: (1) by improving the criteria for faculty hiring, and (2) by providing the faculty with information on what factors are important contributors to improving teaching effectiveness.

LIMITATIONS[1]

The findings presented in this study are limited as follows:

1. Class size was measured by the number of students taking the final examination. An average of the number of students during the whole semester would have been a better measure of class size.

2. No pre-test was available.

---

[1]A new study is currently in progress utilizing pre-tests of student attitudes and achievement, as well as various faculty characteristics including their verbal facility.

3. The sample was limited to those students who had taken the FTGSSC test.

4. Other possible independent variables that could affect student achievement in the final examination, but were not measured, were not included. Among these were: (1) student motivation, (2) the number of hours per week during which the student is gainfully employed, and (3) the number of college semester credits the student is carrying that semester.

It is possible that the inclusion of another independent variable would change the values of the instructor dummy variable coefficients. This would be the case if the added independent variable affected the students of some instructor more than it affected the students of other instructors. (Any evaluation of the sufficiency of the independent variables used to isolate instructor teaching effectiveness should be based on whether or not there are any excluded variables whose effect is not equally distributed among instructors.)

The adequacy of the results is also constrained by the usual limitations of the least-squares multiple regression method. Of the methodological difficulties created by these limitations, simultaneous equation bias, and possible non-linearities were the most serious ones encountered in this study. The first problem, which occurred in connection with class size, could only be corrected with actual experimental controls. The second problem could be corrected by the much simpler technique of trying various non-linear forms of the regression.

Conclusions and Implications

The results indicate that multiple regression with dummy variables for instructors may be used with reservations, discussed above, as a tool to identify the significance of any instructors contribution to student achievement. Significant differences between instructors were identified and it was possible to rank order instructors on their regression coefficients.

As to the importance of the students prior background on his achievement in the course, it was found that his Social Science Florida Twelfth Grade Score explained more variance (approximately 32%) than any other variable. Grade point average, however, and cumulative hours earned do contribute significantly to the explanation of the residual variation.

The classroom condition used, class size, was found to be negatively related to final exam scores but not at the .05 level of significance.

In one case, the department chairman's rank ordering of instructors according to his perception of their contribution to students final exam scores (based on classroom visitations and personal knowledge) matched closely the rank ordering obtained via use of the regression coefficients. In the other situation discussed in this study a chairman less familiar with his staff ranked them differently than the dummy variable coefficients.

It was found that neither salaries paid instructors

nor other pay indices typically used, specifically credits earned and years of teaching experience, were significantly related to the teacher's contribution to student achievement.

Finally, it was observed that while use of all the independent variables explains approximately 48% of the variance, the removal of the class size variable altered the rank order of the instructor's contributions to the final examination score. Since that is still a large portion of unexplained variance, it is suspected that if other significant independent variables could be quantified and included in the analysis, those not equally distributed among instructors might affect the rank order of the instructor's contribution.

In addition to raising questions as to the relationship between salary and performance, the method used in this study suggests itself for possible use as part of the faculty evaluation process. In this respect, it is most important to stress the limitations and qualifications involved in such use, and the way in which it would fit in with other evaluative criteria.

How well an instructor achieves the job of imparting to students previously agreed upon cognitive knowledge is an important part of the teaching job, but as everyone recognizes, only a part. There are many other aspects of the teaching job such as the instructor's contribution to building up a fund of useful knowledge

in his subject. In addition, teaching generally involves more than just transmitting cognitive knowledge. The affective objectives of creating an interest in the subject, imparting an inquisitive attitude, as well as inducing organized work habits, a more open-minded attitude, an ability to work with others, or a multitude of other possible changes in behavior are clearly worthy of pursuit.

Whatever the teaching objectives may be, the important thing is that they be clearly stated and that the weight they carry as evaluation devices correspond to the relative importance given to their achievement. This last condition is usually disregarded, because there is a tendency to overemphasize as evaluation criteria those objectives whose achievement are easily measured, or for which measurable proxy variables exist which are believed to be related to the achievement of an objective. At present, years of teaching experience and graduate credits provide such easily measurable evaluative criteria. A strong emphasis on such a narrow criteria produces an incentive to achieve those measurable objectives, at the expense of others. For example, an instructor may take additional graduate credits because it can earn him a pay raise, but this can keep him too busy studying to do a good job of preparing his lectures.

If a measurement of effectiveness in communicating cognitive knowledge is used as an evaluation device, without also improving upon the measurement of achievement

of other objectives, there exists a danger that it would also be overemphasized as an evaluation device, at the expense of the unmeasurable criteria.

Other issues involved in the use of this method are equity in faculty participation in the composition of the exam and the necessity to provide safeguards against possible tendencies to teach for the test only. The former might be handled by providing for all an equal degree of participation in writing and approving the final examination. The latter possibility may require that for each course a large bank of questions be developed from which a stratified random sample would be chosen.

REFERENCES

1. Astin, A., College Dropouts: A National Profile, Washington, D.C.: American Council on Education, A.C.E. Research Reports, Vol. 7, No. 1, February, 1972.

2. Astin, A. and R. J. Panos, The Educational and Vocational Development of College Students, Washington, D.C.: American Council on Education, 1969.

3. Barro, S., "An Approach to Developing Accountability Measures for the Public Schools." Phi Delta Kappan, 52: 196-205, 1970.

4. Centra, J. A. and D. A. Rock, "College Environments and Student Academic Achievement." American Educational Research Journal. 8: 623-634, 1971.

5. Cohen, A. M. and F. B. Brawer, Measuring Faculty Performance. ERIC ED 031 222, 81 p., 1969.

6. Hanushek, E., The Value of Teachers in Teaching. Santa Monica, California: The RAND Corporation, 1970.